# ESE 568 COMPUTER AND ROBOT VISION

Stony Brook University, Electrical and Computer Engg., Fall 2020, 3 credits.

**Instructor: Prof. Murali Subbarao**

**DRAFT VERSION 1.0. This is subject to some changes.**

Pre-requisites: Basic background in Linear algebra, Claculus, Probability, and Programming. Projects will be in MATLAB. If you have prior programming experience (as in ESE 224), then you will need 8 hours to learn enough MATLAB for this course.

**Text book:**

*Computer Vision: Algorithms and Applications*, Richard Szeliski, Spinger 2010, Available free at **http://szeliski.org/Book/**

**References**

Many online resources.
Some examples:
https://www.cc.gatech.edu/~hays/compvision/
http://vision.stanford.edu/teaching/cs131_fall1617/schedule.html
http://www.cs.cmu.edu/~16385/
http://cs.brown.edu/courses/csci1430/#schedule
http://www.cs.cmu.edu/~16385/s17/
http://www.cs.cmu.edu/afs/cs/academic/class/15385-s12/www/
http://6.869.csail.mit.edu/fa18/materials.html
https://cs.brown.edu/courses/csci1430/proj4/
https://colab.research.google.com/notebooks/welcome.ipynb
http://inst.eecs.berkeley.edu/~cs280/sp15/

**Instructor:** Prof. Murali Subbarao
**Office Hours:** Tue. 11.15 am to 1.15 pm
Thurs. 11.15 am to 1.15 pm

**Part I Image Formation Models and Image Processing**

1. *Introduction:* Introduction, Overview, and applications.
2. Digital images for representing 2D, 3D, and moving objects. Human eye and digital camera models.
3. MATLAB tutorial for computational vision, and Linear algebra overview. (vectors, points, lines, planes, surfaces, matrices). Other CV tools: Python, numpy, OpenCV, Tensor flow, etc.
4. Image recognition paradigm, Quantitative vision for robotics and industry, and qualitative vision for object recognition (e.g. face recognition).
5. *Photometric information: Color:* Physics of color, human perception of color, color models (RGB, HSI).

6. *Geometric-information:* Representation of points, lines, planes, surfaces, and shapes in 3D, nature and structure of medical images. Two-dimensional and three-dimensional geometric transformations of images and 3D scenes.
7. *Image filtering:* gray-level transformations, histograms, convolution, noise reduction, spatial and Fourier domain filtering and convolution, Gaussian filtering, and image resolution pyramids.

**Part II Image Features: detection and matching**

8. *Feature detection*: gradient vector, Canny's edge detection, Harris-corner detector.

## <u>Mid-term test 1.</u>

9. *Contours:* Model fitting, Total LSE, Least Median Square Error.
10. RANSAC, Hough transform.
11. SIFT vector, image stitching.
12. *Pattern classification and Image segmentation:* Image features, SIFT and related feature vectors, clustering techniques, K-mean clustering. PCA.

**Part III 3D Imaging, 3D Motion, Medical imaging.**

13. *Three-dimensional shape recovery:* 3D from Stereo Images; Stereo Camera model, calibration, matching, rectification.
14. structured-light, RGBD cameras, Laser and LIDAR, and related techniques.
15. 3D Motion from Video, optical flow, other shape-from-x methods (texture, shading, focus/defocus, Optical flow, etc). Machine and robot vision applications and self-driving cars.
16. *Medical Imaging:* Modes of medical imaging, X-ray Computed Tomography, image reconstruction algorithms.

## <u>Mid-term test 2.</u>

**Part IV High-level Vision: Machine Learning, Neural Nets, and Artificial Intelligence**

17. Machine learning principles and techniques for object recognition. Nearest-neighbor, nearest centroid, K-NN.
18. Support Vector Machines.
19. Neural Nets, Convolution Neural Nets,
20. Deep learning, AI.

**Final Quiz (10%. Final exam will be a 30 minute quiz, with questions having short answers).**

Programming Projects (30%) : There will be around 3 programming projects using MATLAB. Each project may take around 10 hours for completion.

Project 1: 2D and 3D Geometric transforms, imaging in a pin-hole camera.
Project 2: Image processing, Feature Detection, and Local Feature Descriptor
Project 3: Convolutional Neural Nets for Image Recognition
**Seminar presentation (10%):** Each student will have to present a paper published within the last 10 years on a topic of current interest. Length of presentation: 20 minutes.

**GRADING**

| | | | |
|---|---|---|---|
| **Mid-term Test 1** | **:** | **25%** | **(1 hr 20 mins)** |
| **Mid-term Test 2** | **:** | **25%** | **(1 hr 20 mins)** |
| **Final Quiz** | **:** | **10%** | **(30 mins)** |
| **Projects:** | **:** | **30%** | |
| **Presentation** | **:** | **10%** | |

**Grading Policy**
Grades are assigned based on absolute percentage of total marks as below.
A : 93—100 , A- : 88—92 ,
B+ : 83—87, B : 78—82, B- : 73--77
C+ : 70—72, C : 65—69, C- : 61—64,
D+ : 56—60, D : 51—55, F : 0—50